Phenotypic Characterization of Microorganisms by Barcoded Transposon Mutagenesis

Adam Deutschbauer[1,4], Julia Oh[2], Morgan Price[1,4], Paramvir Dehal[1,4]* (PSDehal@lbl.gov), Dan Bruno[2], Marlene Henriquez[1,4], Romy Chakraborty[1,4], Terry C. Hazen[1,4], Corey Nislow[3], Guri Giaever[3], Ronald W. Davis[2], Adam P. Arkin[1,4]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Stanford Genome Technology Center, Stanford, CA; [3]University of Toronto, Toronto, ON, Canada; [4]VIMSS (Virtual Institute of Microbial Stress and Survival) http://vimss.lbl.gov/

ESPP Environmental Stress Pathway Project

DOE GENOMICS:GTL ACCELERATING DISCOVERY FOR ENERGY AND ENVIRONMENT

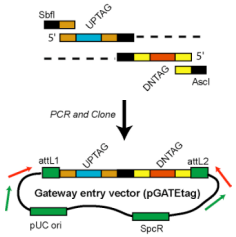
VIMSS Virtual Institute for Microbial Stress and Survival | Miami University | University of Missouri | University of Washington | Diversa | Sandia National Laboratories | Oak Ridge National Laboratory | University of Oklahoma | Berkeley University of California | 75 Years of World Class Science

http://vimss.lbl.gov/

Introduction

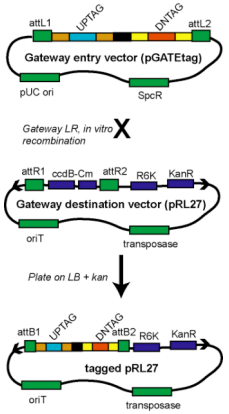
The ability to sequence genomes far outpaces our ability to systematically determine gene function. Subsequently, systems-level analyses of less studied bacteria are limited by the presence of numerous uncharacterized genes and an over reliance on annotations from well studied bacteria such as *E. coli*. To meet this challenge, we are developing a flexible mutagenesis technique and applying it to the environmental bacteria *Shewanella oneidensis* MR1, *Shewanella loihica* PV4, and *Desulfovibrio desulfuricans* G20. The hallmarks of our approach are the sequencing and archiving of thousands of transposon mutants and the use of molecular "barcodes" (tags) for the parallel phenotypic analysis of defined mutant pools. The successful completion of this project will enable the quantitative phenotypic analysis of thousands of mutants across a wide range of conditions. In addition, our genetic resources provide a framework for the systematic genetic interrogation of individual pathways.

Cloning of Tag Modules



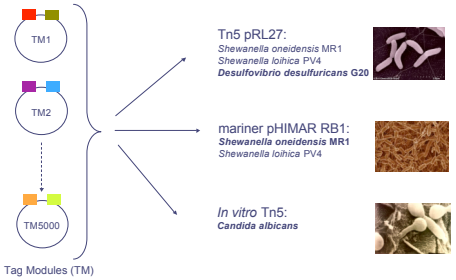
The tag modules are constructed using two long oligonucleotides. one oligo contains the UPTAG, a second oligo contains the DOWNTAG. The tags are identical to those used in the yeast deletion system. Each tag module is cloned into pGATEtag, a modified Gateway entry vector. To date, we have cloned and sequence-verified 3000 unique tag modules. Our aim is to construct ~5000 unique tag modules so that the majority of nonessential genes can be represented in a single mutant pool.

Transfer of Tag Modules via Gateway Reaction



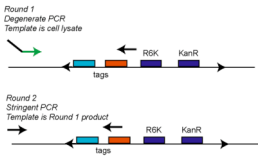
The tag modules are universal. They can be transferred *in vitro* to any Gateway compatible destination vector (either as single reactions or in pools). We converted the Tn5-transposon delivery plasmid pRL27 into a Gateway entry vector by cloning the "AttR1-ccdB-CmR-AttR2" cassette into a KpnI site located at the end of Tn5.

Our Focus Organisms



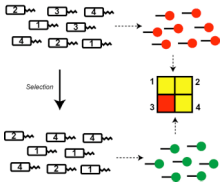
We have converted pRL27 (Tn5), pHIMAR-RB1 (mariner), and an *in vitro* Tn5 system (epicentre) into Gateway destination vectors. The tagged transposons are functional in the indicated organisms.

High-throughput Transposon Insertion Mapping



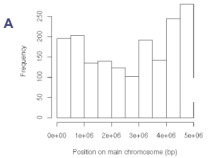
- After random mutagenesis with the collection of tagged transposons, we determine the insertion location of the transposon in the genome and the tag identify using a 2-step degenerate PCR approach. Sequence analysis of the second round PCR product is used to infer both the tag identify and the genome insertion site.
- We anticipate sequencing ~30,000 colonies per bacterial genome in order to get adequate coverage. Mutants with unique insertion sites will be selected from these libraries and combined into pools of ~5000 mutants (each with a unique tag module).
- To facilitate this process, we have optimized the use of a liquid handling robot (Biomek FX) to process (PCR and sequencing) 1536 mutants per day.

Parallel Analysis of Mutant Pools

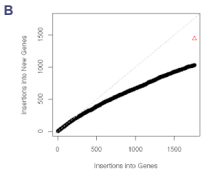


The basic premise behind the parallel phenotypic analysis of mutant pools is illustrated for a single pool of only four mutants (marked 1 to 4). Both prior to and post-selection, genomic DNA is extracted from the pool and all of the UPTAGs and DOWNTAGs are PCR amplified in two separate reactions (indicated here for only one of the two tags). The amplified barcodes from before and after selection, labeled with two different fluorescent dyes, are combined and hybridized to an Affymetrix microarray containing the tag complements. In this example, mutant3 exhibited a fitness defect. With our approach, we can monitor the fitness of 5000 mutants in parallel.

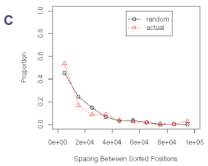
Application to *Shewanella oneidensis* MR1



A. In a pilot study, ~3200 *Shewanella* tagged Tn5 mutants were sequenced. We determined the tag identity and insertion location in ~67% of the mutants. Shown is the insertion location for 1937 mutants on main chromosome. There is a bias towards insertions near the origin of replication and away from the terminus.



B. Displayed is the increase in new genes hit with the transposon as a function of mapped insertion sites. The number of new genes hit falls short of expectations (red triangle). Given that Tn5 may exhibit some insertion biases, we have adapted a mariner-based transposon for use with *Shewanella*.



C. The spacing between neighboring insertion sites among 1937 mutants indicates a small bias for tightly spaced insertions. These "hot spots" argue for the use of a second, independent transposon for constructing the library.

Future Directions

- Generate ~2000 additional sequence-defined tag modules.
- Go into production phase and rapidly generate thousands of sequence-defined transposon mutants in multiple bacterial genomes.
- Explore the potential for double deletion construction for global genetic interaction studies.
- Automate and miniaturize mutant pool experiments using procedures developed for the yeast deletion collection. Such an effort will enable us to profile comprehensive mutant libraries across hundreds of diverse conditions.

ACKNOWLEDGEMENT

ESPP is part of the Virtual Institute for Microbial Stress and Survival supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics Program:GTL through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.